

Extracting drug indications and adverse drug reactions from Spanish health social media

Isabel Segura-Bedmar, Santiago de la Peña, Paloma Martínez

Computer Science Department

Carlos III University of Madrid, Spain

{isegura|spena|pmf}@inf.uc3m.es

Abstract

In this paper, we present preliminary results obtained using a system based on co-occurrence of drug-effect pairs as a first step in the study of detecting adverse drug reactions and drug indications from social media texts. To the best of our knowledge, this is the first work that extracts this kind of relationships from user messages that were collected from an online Spanish health-forum. In addition, we also describe the automatic construction of the first Spanish database for drug indications and adverse drug reactions.

1 Introduction

The activity of Pharmacovigilance (science devoted to the detection and prevention of any possible drug-related problem, including adverse drug effects) has gained significant importance in the recent decades, due to the growing number of drug safety incidents (Bond and Raehl, 2006) as well as to their high associated costs (van Der Hooft et al., 2006).

Nowadays, the major medicine regulatory agencies such as the US Food and Drug Administration (FDA) or the European Medicines Agency (EMA) are working to create policies and practices to facilitate the reporting of adverse drug reactions (ADRs) by healthcare professionals and patients. However, several studies have shown that ADRs are under-estimated because many healthcare professionals do not have enough time to use the ADR reporting systems (Bates et al., 2003; van Der Hooft et al., 2006; McClellan, 2007). In addition, healthcare professionals tend to report only those ADRs on which they have absolute certainty of their existence. Unlike reports from healthcare professionals, patient reports often provide more detailed and explicit information

about ADRs (Herxheimer et al., 2010). Nevertheless, the rate of ADRs reported by patients is still very low probably because many patients are still unaware of the existence of ADR reporting systems. In addition, patients may feel embarrassed when describing their symptoms.

In this paper, we pose the hypothesis that health-related social media can be used as a complementary data source to the ADR reporting systems. In particular, health forums contain a large number of comments describing patient experiences that would be a fertile source of data to detect unknown ADRs.

Several systems have been developed for extracting ADRs from social media (Leaman et al., 2010; Nikfarjam and Gonzalez, 2011). However to the best of our knowledge, only one work in the literature has focused on the detection of ADRs from social media in Spanish (Segura-Bedmar et al., 2014). Indeed, it is only concerned with the detection of mentions of drugs and their effects, without dealing with the extraction of the relationships between them. In this paper, we extend this existing work in order to extract drug indications and adverse drug reactions from user comments in a Spanish health-forum.

The remaining of this paper is structured as follows: the next section surveys related work on ADR detection from social media. Section 3 describes the creation of a gold-standard corpus we used for our experiments. Sections 4 and 5 respectively describe the techniques employed and their results. Lastly, some conclusive remarks and future perspectives are given in Section 6.

2 Related Work

In recent years, the application of Natural Language Processing (NLP) techniques to mine drug indications and adverse drug reactions from texts has been explored with promising results, mainly in the context of drug labels (Gurulingappa et al.,

2013; Li et al., 2013; Kuhn et al., 2010; Fung et al., 2013), biomedical literature (Xu and Wang, 2013), medical case reports (Gurulingappa et al., 2012) and health records (Friedman, 2009; Sohn et al., 2011). However, as it will be described below, the extraction of these drug relationships from social media has received much less attention.

To date, most of research on drug name recognition concerns either biomedical literature (Segura-Bedmar et al., 2013; Krallinger et al., 2013) or clinical records (Uzuner et al., 2010), thus leaving unexplored this task in social media texts.

To our knowledge, there is no work in the literature that addresses the extraction of drug indications from social media texts. Regarding the detection of ADRs, Leaman et al., (2010) developed a system to automatically recognize adverse effects in user comments from the DailyStrength¹ health-related social network. A corpus of 3,600 comments was manually annotated with a total of 1,866 drug conditions, including beneficial effects, adverse effects, indications and others. This study focused only on a set of four drugs, and thereby, drug name recognition was not addressed. The system used a dictionary-based approach to identify adverse effects and a set of keywords in order to distinguish adverse effects from the other drug conditions. The dictionary consisted of 4,201 concepts, which were collected from several resources such as the COSTART vocabulary (FDA, 1970), the SIDER database (Kuhn et al., 2010), the MedEffect database² and a list of colloquial phrases manually collected from the comments. The system achieved a precision of 78.3% and a recall of 69.9% (an f-measure of 73.9%).

Later, Nikfarjam and Gonzalez (2011) applied association rule mining to extract frequent patterns describing opinions about drugs. The rules were generated using the Apriori tool, an implementation of the Apriori algorithm (Agrawal et al., 1994) for association rule mining. The main advantage of this approach over the dictionary based approach is that the system is able to detect terms not included in the dictionary. The results of this study were 70.01% precision and 66.32% recall, for an f-measure of 67.96%.

Benton et al.,(2011) collected a lexicon of lay medical terms from websites and databases about drugs and their adverse effects to identify drug ef-

fects. Then, the authors applied the Fishers exact test (Fisher, 1922) to find all the drug-effect pairs that co-occurred independently by chance in a corpus of user comments. To evaluate the system, the authors focused only on the four most commonly used drugs to treat breast cancer. Precision and recall were calculated by comparing the adverse effects from their drug labels and the adverse effects obtained by the system. The system obtained an average precision of 77% and an average recall of 35.1% for all four drugs.

To the best of our knowledge, the system described in (Segura-Bedmar et al., 2014) is the only one that has dealt with the detection of drugs and their effects from Spanish social media streams. The system used the Textalytics tool³, which follows a dictionary-based approach to identify entities in texts. The dictionary was constructed based on the following resources: CIMA⁴ and MedDRA⁵. CIMA is an online information center maintained by the Spanish Agency for Medicines and Health Products (AEMPS). CIMA provides information on all drugs authorized in Spain, though it does not include drugs approved only in Latin America. CIMA contains a total of 16,418 brand drugs and 2,228 generic drugs. Many brand drugs have very long names because they include additional information such as dosages, mode and route of administration, laboratory, among others (for example, *ESPIDIFEN 400 mg GRANULADO PARA SOLUCION ORAL SABOR ALBARI-COQUE*). For this reason, brand drug names were simplified before being included in the dictionary. After removing the additional information, the resulting list of brand drug names consisted of 3,662 terms. Thus, the dictionary contained a total of 5,890 drugs. As regards to the effects, the authors decided to use MedDRA, a medical multilingual terminology dictionary about events associated with drugs. MedDRA is composed of a five levels hierarchy. A total of 72,072 terms from the most specific level, "Lowest Level Terms" (LLTs), were integrated into the dictionary. In addition, several gazetteers including drugs and effects were collected from websites such as Vademecum⁶, a Spanish online website that provides information to patients on drugs and their side effects, and

¹<http://www.dailystrength.org/>

²<http://www.hc-sc.gc.ca/dhp-mps/medeff/index-eng.php>

³<https://textalytics.com/>

⁴<http://www.aemps.gob.es/cima/>

⁵<http://www.meddra.org/>

⁶<http://www.vademecum.es/>

the ATC system⁷, a classification system of drugs. Thus, the dictionary and the two gazetteers contained a total of 7,593 drugs and 74,865 effects. The system yielded a precision of 87% for drugs and 85% for effects, and a recall of 80% for drugs and 56% for effects.

3 The SpanishADR corpus

Segura-Bedmar et al., (2014) created the first Spanish corpus of user comments annotated with drugs and their effects. The corpus consists of 400 comments, which were gathered from ForumClinic⁸, an interactive health social platform, where patients exchange information about their diseases and their treatments. The texts were manually annotated by two annotators with expertise in Pharmacovigilance. All the mentions of drugs and effects were annotated, even those containing spelling or grammatical errors (for example, *hemorragia* (haemorrhage)). An assessment of the inter-annotator agreement (IAA) was based on the F-measure metric, which approximates the kappa coefficient (Cohen, 1960) when the number of true negatives (TN) is very large (Hripcsak and Rothschild, 2005). This assessment revealed that while drugs showed a high IAA (0.89), their effects point to moderate agreement (0.59). This may be due to drugs have specific names and there are a limited number of them, however their effects are expressed by patients in many different ways due to the variability and richness of natural language. The corpus is available for academic purposes⁹.

In this paper, we extend the Spanish corpus to incorporate the annotation of the relationships between drugs and their effects. In particular, we annotated drug indications and adverse drug reactions. These relationships were annotated at comment level rather than sentence level, because determining sentence boundaries in this kind of texts can be problematic since many users often write ungrammatical sentences. Guidelines were created by two annotators (A1, A2) and a third annotator (A3) was trained on the annotation guidelines. Then, we split the corpus in three subsets, and each subset was annotated by one annotator. Finally, IAA was measured using kappa-statistic on a sample of 97 documents randomly selected. These documents were annotated by the three an-

notators and annotation differences were analysed.

As Table 1 shows, the resulting corpus has 61 drug indications and 103 adverse drug reactions. The average size of a comment is 72 tokens. The average size of a text fragment describing a drug indication is 34.7 tokens and 28.2 tokens for adverse drug reactions.

Annotation	Size
drugs	188
effect	545
drug indication	61
adverse drug reaction	103

Table 1: Size of the extended SpanishADR corpus.

As it is shown in Table 2, the IAA figures clearly suggest that the annotators have high agreement among them. We think that the IAA figures were lower with the third annotator because he did not participate in the guidelines development process, and maybe, he was not trained well enough to perform the task. The main source of disagreement among the annotators could arise from considering whether a term refers to a drug effect or not. This is due to some terms are too general (such as *trastorno* (upset), *enfermedad* (disease), *molestia* (ache)). The annotators A1 and A2, in general, ruled out all the relation instances where these general terms occur, however they were considered and annotated by the third annotator.

	A2	A3
A1	0.8	0.69
A2	-	0.68

Table 2: Pairwise IAA for each combination of two annotators. IAA was measured using Cohens' kappa statistic

4 Methods

In this contribution, some refinements to the system (Segura-Bedmar et al., 2014) are proposed. The error analysis performed in (Segura-Bedmar et al., 2014) showed that most of false positives for drug effects were mainly due to the inclusion of MedDRA terms referring to procedures and tests in the dictionary. MedDRA includes terms for diseases, signs, abnormalities, procedures and tests. Therefore, we decided not to include terms corresponding to the "Procedimientos

⁷http://www.whocc.no/atc_ddd_index/

⁸<http://www.forumclinic.org/>

⁹<http://labda.inf.uc3m.es/SpanishADRCorpus>

médicos y quirúrgicos” and ”Exploraciones complementarias” categories since they do not represent drug effects. Thus, we created a new dictionary that only includes those terms from MedDRA that actually refer to drug effects. As in the system (Segura-Bedmar et al., 2014), we applied the Textalytics tool, which follows a dictionary-based approach, to identify drugs and their effects occurring in the messages. We created a GATE¹⁰ pipeline application integrating the Textalytic module and the gazetteers collected from the Vademecum website and the ATC system proposed in (Segura-Bedmar et al., 2014).

In addition, we created an additional gazetteer in order to increase the coverage. We developed a web crawler to browse and download pages related to drugs from the MedLinePlus website¹¹. Unlike Vademecum, which only contains information for drugs approved in Spain, MedLinePlus also includes information about drugs only approved in Latin America. Terms describing drug effects were extracted by regular expressions from these pages and then were incorporated into a gazetteer. Then, the new gazetteer was also integrated into the GATE pipeline application to identify drugs and effects. Several experiments with different settings of this pipeline are described in the following section.

The main contribution of this paper is to propose an approach for detecting relationships between drugs and their effects from user comments in Spanish. The main difficulty in this task is that although there are several English databases such as SIDER or MedEffect with information about drugs and their side effects, none of them are available for Spanish. Moreover, these resources do not include drug indications. Thus, we have automatically built the first database, *SpanishDrug-EffectBD*, with information about drugs, their drug indications as well as their adverse drug reactions in Spanish. Our first step was to populate the database with all drugs and effects from our dictionary. Figure 1 shows the database schema. Active ingredients are saved into the *Drug* table, and their synonyms and brand names into the *DrugSynset* table. Likewise, concepts from MedDRA are saved into the *Effect* table and their synonyms are saved into the *EffectSynset* table. As it is shown in Figure 1, the database is also de-

signed to store external ids from other databases. Thus, drugs and effects can be linked to external databases by the tables *has_externalIDDDrug* and *has_externalIDDEffect*, respectively.

To obtain the relationships between drugs and their effects, we developed several web crawlers in order to gather sections describing drug indications and adverse drug reactions from drug package leaflets contained in the following websites: MedLinePlus, Prospectos.Net¹² and Prospectos.org¹³. Once these sections were downloaded, their texts were processed using the TextAlyticis tool to recognize drugs and their effects. As each section (describing drug indications or adverse drug effects) is linked to one drug, we decided to consider the effects contained in the section as possible relationships with this drug. The type of relationship depends on the type of section: drug indication or adverse drug reaction. Thus for example, a pair (drug, effect) from a section describing drug indications is saved into the *DrugEffect* table as a drug indication relationship, while if the pair is obtained from a section describing adverse drug reactions, then it is saved as an adverse drug reaction. This database can be used to automatically identify drug indications and adverse drug reactions from texts. Table 3 shows the number of drugs, effects and their relationships stored into the database.

	Concepts	Synonyms
drugs	3,244	7,378
effects	16,940	52,199
drug indications	4,877	
adverse drug reactions	58,633	

Table 3: Number of drugs, effects, drug indications and adverse drug effects in the *SpanishDrug-EffectBD* database.

As regards to the extraction of the relationships between drugs and their effects occurring in the corpus, first of all, texts were automatically annotated with drugs and effects using the GATE pipeline application. Then, in order to generate all possible relation instances between drugs and their effects, we considered several sizes of window: 10, 20, 30, 40 and 50. Given a size n , any pair (drug, effect) co-occurring within a window of n -tokens are treated as a relation instance. Af-

¹⁰<http://gate.ac.uk/>

¹¹<http://www.nlm.nih.gov/medlineplus/spanish/>

¹²<http://www.prospectos.net/>

¹³<http://prospectos.org/>

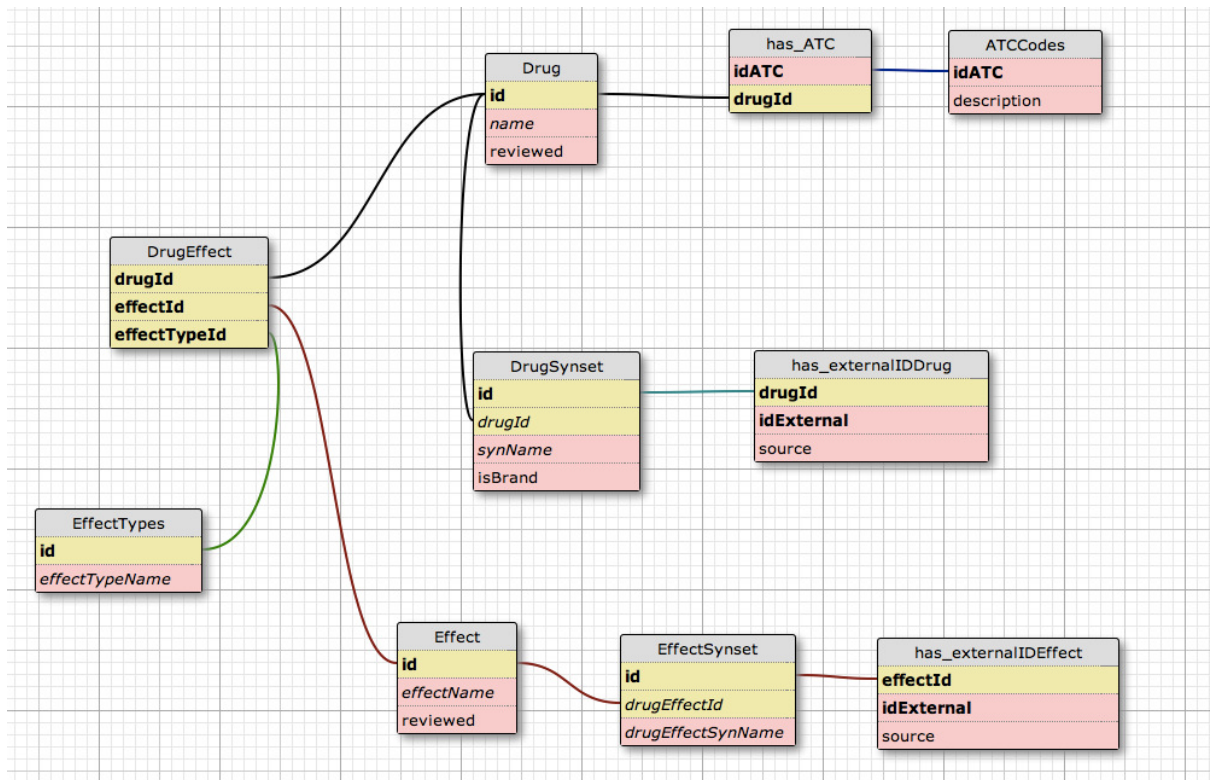


Figure 1: The SpanishDrugEffectBD database schema

terwards, each relation instance is looked up in the *DrugEffect* table in order to determine if it is a positive instance and if this is the case, its type: drug indication or adverse drug reaction.

5 Experiments

Several experiments have been performed in order to evaluate the contribution of the proposed methods and resources. Table 4 shows the results for the named entity recognition task of drugs and effects using the dictionary integrated into the TextAlytic tool. The first row shows the results with the dictionary built from the CIMA and MedDRA resources, while the second one shows the results obtained using the new dictionary in which those MedDRA terms corresponding to "Procedimientos médicos y quirúrgicos" and "Exploraciones complementarias" categories were ruled out. As it can be seen in this table, the new dictionary permits to obtain a significant improvement with respect to the original dictionary. For effect type, precision was increased almost a 40% and recall a 7%. As regards to the contribution of the gazetteers, the coverage for effects improves almost a 6% but with significant decrease in precision of almost 21%. Regarding to the detection of

drugs, the use of gazetteers improves slightly the precision and achieves a significant improvement in the recall of almost 35%.

The major cause of false negatives for drug effects was the use of colloquial expressions (such as *'me deja ko'* (it makes me ko)) to describe an adverse effect. These phrases are not included in our dictionary. Another important cause was the dictionary and gazetteers do not cover all the lexical variations of a same effect (for example *depresión* (depression), *depresivo* (depress), *me deprimo* (I get depressed)). In addition, many false negatives were due to spelling mistakes (for example *hemorragia* instead of *hemorragia* (haemorrhage)) and abbreviations (*depre* is an abbreviation for *depresión* (depression)).

Regarding to the results for the relation extraction task, Table 5 shows the overall results obtained using a baseline system, which considers all pairs (drug, effect) occurring in messages as positive relation instances, and a second approach using the SpanishDrugEffectBD database (a relation instance is positive only if it is found into the database). In both experiments, a window size of 250 tokens was used. The database provides a high precision but with a very low recall of only 15%.

Approach	Entity	P	R	F1
Dictionary	drugs	0.84	0.46	0.60
	effect	0.45	0.38	0.41
New dictionary	drugs	0.84	0.46	0.60
	effect	0.84	0.45	0.59
New dictionary plus gazetteers	drugs	0.86	0.81	0.84
	effect	0.63	0.51	0.57

Table 4: Precision, Recall and F-measure for named entity recognition task.

As it can be seen in Table 6, when the type of the relationship is considered, the performance is even lower.

Approach	P	R	F1
Baseline	0.31	1.00	0.47
SpanishDrugEffectBD	0.83	0.15	0.25

Table 5: Overall results for relation extraction task (window size of 250 tokens).

Relation	P	R	F1
Drug indication	0.50	0.02	0.03
Adverse drug reaction	0.65	0.11	0.18

Table 6: Results for drug indications and adverse drug reactions using only the database (window size of 50 tokens).

Figure 2 shows an example of the output of our system using the database. The system is able to detect the relationship of indication between *alprazolman* and *ansiedad* (anxiety), but fails in detecting the adverse drug reaction between *alprazolman* and *dependencia* (dependency). The adverse drug reaction between *lamotrigina* and *vertigo* is detected.

The co-occurrence approach provides better results than the use of the database. Table 7 shows the results for different size of windows. As it was expected, small sizes provide better precision but lower recall.

6 Conclusion

In this paper we present the first corpus where 400 user messages from a Spanish health social network have been annotated with drug indications and adverse drug reactions. In addition, we present preliminary results obtained using a very simple system based on co-occurrence of drug-effect pairs as a first step in the study of detecting

Size of window	P	R	F1
10	0.71	0.24	0.36
20	0.59	0.53	0.56
30	0.52	0.69	0.59
40	0.47	0.77	0.58
50	0.44	0.84	0.58

Table 7: Overall results for relation extraction task using the co-occurrence approach considering different window sizes.

adverse drug reactions and drug indications from social media streams. Results show that there is still much room for improvement in the identification of drugs and effects, as well as in the extraction of drug indications and adverse drug reactions.

As it was already mentioned in Section 2, the recognition of drugs in social media texts has hardly been tackled since most systems were focused on a given and fixed set of drugs. Moreover, little research has been conducted to extract relationships between drugs and their effects from social media. Most systems for extracting ADRs follow a dictionary-based approach. The main drawback of these systems is that they fail to recognize terms which are not included in the dictionary. In addition, the dictionary-based approach is not able to handle the large number of spelling and grammar errors in social media texts. Moreover, the detection of ADRs and drug indications has not been attempted for languages other than English. Indeed, automatic information extraction from Spanish-language social media in the field of health remains largely unexplored.

Social media texts pose additional challenges to those associated with the processing of clinical records and medical literature. These new challenges include the management of meta-information included in the text (for example as tags in tweets)(Bouillot et al., 2013), the detection of typos and unconventional spelling, word short-

Results - Drugs: Indications and Adverse Effects

Text Annotated

DEPR357 segun muchos prospectos no se deben pasar 21 dias (3 semanas). . el alprazolam es precisamente el ansiolitico que mas dependencia genera por su alta potencia y vida media corta. . pero si tuviera que darte mi opinion, te diria que depende de la "adiccion a la ansiedad" que tengas. si andas medio tranquilo y bien, la dependencia no es tanta y los puedes soltar con facilidad; si tu ansiedad es altisima no los vas a poder dejar. . un saludo vecino

TRBI1122 hola anais. . finalmente conseguí hablar con mi psiquiatra habitual y me dijo que el vertigo era un efecto secundario de la lamotrigina. le comente que lo habia suspendido. asi esta informado de mi situacion

Drug Effect Indication AdverseEffect

Figure 2: An example of the output of the system using the database.

enings (Neunerdt et al., 2013; Moreira et al., 2013) and slang and emoticons (Balahur, 2013), among others. Another challenge that should be taken into account is that while clinical records and medical literature can be mapped to terminological resources or biomedical ontologies, lay terminology used by patients to describe their treatments and their effects, in general, is not collected in any terminological resource, which would facilitate the automatic processing of this kind of texts.

In this paper, we also describe the automatic creation of a database for drug indications and adverse drug reactions from drug package leaflets. To the best of our knowledge, this is the first database available for Spanish. Although the use of this database did not improve the results due to its limited coverage, we think that the database could be a valuable resource for future efforts. Thus, we plan to translate the database into an ontology and to populate it with more entities and relationships. As future work, we plan the following tasks:

- To create a lexicon containing idiomatic expressions used by patients to express drug effects.
- To use techniques such as lemmatization and stemming to cope with the problem of lexical variability and to resolve abbreviations.
- To integrate advanced matching methods capable of dealing with the spelling error problem.
- To increase the size of the corpus.

- To apply a SVM classification approach to extract relationships between drugs and their effects.

We hope our research will be beneficial to AEMPS as well as to the pharmaceutical industry in the improvement of their pharmacovigilance systems. Both the corpus and the database are freely available online¹⁴ for research purposes.

Acknowledgments

This work was supported by the EU project Trend-Miner [FP7-ICT287863], by the project MULTIMEDICA [TIN2010-20644-C03-01], and by the Research Network MA2VICMR [S2009/TIC-1542].

References

- Rakesh Agrawal, Ramakrishnan Srikant, et al. 1994. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499.
- Alexandra Balahur. 2013. Sentiment analysis in social media texts. *WASSA 2013*, page 120.
- David W Bates, R Scott Evans, Harvey Murff, Peter D Stetson, Lisa Pizziferri, and George Hripcsak. 2003. Detecting adverse events using information technology. *Journal of the American Medical Informatics Association*, 10(2):115–128.
- Adrian Benton, Lyle Ungar, Shawndra Hill, Sean Hennessy, Jun Mao, Annie Chung, Charles E Leonard, and John H Holmes. 2011. Identifying potential adverse effects using the web: A new approach to

¹⁴<http://labda.inf.uc3m.es/SpanishADRCorpus>

- medical hypothesis generation. *Journal of biomedical informatics*, 44(6):989–996.
- CA Bond and Cynthia L Raehl. 2006. Adverse drug reactions in united states hospitals. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 26(5):601–608.
- Flavien Bouillot, Phan Nhat Hai, Nicolas Béchet, Sandra Bringay, Dino Ienco, Stan Matwin, Pascal Poncelet, Mathieu Roche, and Maguelonne Teisseire. 2013. How to extract relevant knowledge from tweets? In *Information Search, Integration and Personalization*, pages 111–120. Springer.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- FDA. 1970. National adverse drug reaction directory: Costart (coding symbols for thesaurus of adverse reaction terms). *Rock-Irvine, Charles F, Sharp, r, MD, Huntington Memorial Hospital, Stuart I, Silverman, MD, University of California, Los Angeles, West Los Angeles-Veterans Affairs Medical Center, Osteoporosis Medical Center*.
- Ronald A Fisher. 1922. On the interpretation of chi-squared from contingency tables, and the calculation of p. *Journal of the Royal Statistical Society*, 85(1):87–94.
- Carol Friedman. 2009. Discovering novel adverse drug events using natural language processing and mining of the electronic health record. In *Artificial Intelligence in Medicine*, pages 1–5. Springer.
- Kin Wah Fung, Chiang S Jao, and Dina Demner-Fushman. 2013. Extracting drug indication information from structured product labels using natural language processing. *Journal of the American Medical Informatics Association*, 20(3):482–488.
- Harsha Gurulingappa, Abdul Mateen-Rajput, Luca Toldo, et al. 2012. Extraction of potential adverse drug events from medical case reports. *J Biomed Semantics*, 3(1):15.
- Harsha Gurulingappa, Luca Toldo, Abdul Mateen Rajput, Jan A Kors, Adel Taweel, and Yorki Tayrouz. 2013. Automatic detection of adverse events to predict drug label changes using text and data mining techniques. *Pharmacoepidemiology and drug safety*, 22(11):1189–1194.
- A Herxheimer, MR Crombag, and TL Alves. 2010. Direct patient reporting of adverse drug reactions. a twelve-country survey & literature review. *Health Action International (HAI)(Europe)*. Amsterdam.
- George Hripcsak and Adam S Rothschild. 2005. Agreement, the f-measure, and reliability in information retrieval. *Journal of the American Medical Informatics Association*, 12(3):296–298.
- Martin Krallinger, Florian Leitner, Obdulia Rabal, Miguel Vazquez, Julen Oyarzabal, and Alfonso Valencia. 2013. Overview of the chemical compound and drug name recognition (chemdner) task. In *BioCreative Challenge Evaluation Workshop vol. 2*, page 2.
- Michael Kuhn, Monica Campillos, Ivica Letunic, Lars Juhl Jensen, and Peer Bork. 2010. A side effect resource to capture phenotypic effects of drugs. *Molecular systems biology*, 6(1).
- Robert Leaman, Laura Wojtulewicz, Ryan Sullivan, Annie Skariah, Jian Yang, and Graciela Gonzalez. 2010. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In *Proceedings of the 2010 workshop on biomedical natural language processing*, pages 117–125. Association for Computational Linguistics.
- Qi Li, Louise Deleger, Todd Lingren, Haijun Zhai, Megan Kaiser, Laura Stoutenborough, Anil G Jegga, Kevin Bretonnel Cohen, and Imre Solti. 2013. Mining fda drug labels for medical conditions. *BMC medical informatics and decision making*, 13(1):53.
- Mark McClellan. 2007. Drug safety reform at the fdapendulum swing or systematic improvement? *New England Journal of Medicine*, 356(17):1700–1702.
- Silvio Moreira, Joao Filgueiras, and Bruno Martins. 2013. Reaction: A naive machine learning approach for sentiment classification. In *Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013)*, page 490.
- Melanie Neunerdt, Michael Reyer, and Rudolf Mathar. 2013. A pos tagger for social media texts trained on web comments. *Polibits*, 48:59–66.
- Azadeh Nikfarjam and Graciela H Gonzalez. 2011. Pattern mining for extraction of mentions of adverse drug reactions from user comments. In *AMIA Annual Symposium Proceedings*, volume 2011, page 1019. American Medical Informatics Association.
- Isabel Segura-Bedmar, Paloma Martínez, and María Herrero-Zazo. 2013. Semeval-2013 task 9: Extraction of drug-drug interactions from biomedical texts (ddiextraction 2013). *Proceedings of Semeval*, pages 341–350.
- Isabel Segura-Bedmar, Ricardo Revert, and Paloma Martnez. 2014. Detecting drugs and adverse events from spanish social media streams. In *Proceedings of the 5th International Louhi Workshop on Health Document Text Mining and Information Analysis (Louhi 2014)*.
- Sunghwan Sohn, Jean-Pierre A Kocher, Christopher G Chute, and Guergana K Savova. 2011. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. *Journal of the American Medical Informatics Association*, 18(Suppl 1):i144–i149.

- Özlem Uzuner, Imre Solti, and Eithon Cadag. 2010. Extracting medication information from clinical text. *Journal of the American Medical Informatics Association*, 17(5):514–518.
- Cornelis S van Der Hooft, Miriam CJM Sturkenboom, Kees van Grootheest, Herre J Kingma, and Bruno H Ch Stricker. 2006. Adverse drug reaction-related hospitalisations. *Drug Safety*, 29(2):161–168.
- Rong Xu and QuanQiu Wang. 2013. Large-scale extraction of accurate drug-disease treatment pairs from biomedical literature for drug repurposing. *BMC bioinformatics*, 14(1):181.